

## **Hypothetical Identities: Explanatory Problems for the Explanatory Argument**

*Philosophical Psychology* 2013

Markus I. Eronen

Post-doctoral Researcher

Ruhr University Bochum

Institute for Philosophy II; GA03/150

Universitätsstraße 150

D-44780 Bochum

Germany

Email: Markus.Eronen@rub.de

**Abstract:** Recently several philosophers have defended an explanatory argument that supposedly provides novel empirical grounds for accepting the type identity theory of phenomenal consciousness. They claim that we are justified in believing that the type identity thesis is true because it provides the best explanation for the correlations between physical properties and phenomenal properties. In this paper, I examine the actual role identities play in science and point out crucial shortcomings in the explanatory argument. I show that the supporters of the argument have failed to show that the identity thesis provides a satisfactory explanation for the correlations between physical and phenomenal properties. Hence, the explanatory argument, as it stands, does not provide new grounds for accepting the type identity theory.

### **1. Introduction**

The relation of mental properties to physical properties has preoccupied philosophers throughout the history of modern philosophy of mind. Particularly problematic are

phenomenal properties or qualia, which apparently resist physical explanations. Type physicalism (also known as type identity theory, type materialism or reductive physicalism) is the view that mental properties, including phenomenal properties, are identical to physical properties. Recently several philosophers (Bates, 2009; Block & Stalnaker, 1999; Hill, 1991; McLaughlin, 2001, 2007, 2010; see also Melnyk, 2003) have advanced new *positive* grounds for accepting type physicalism, arguing that it provides the *best explanation* for the correlations between phenomenal and physical properties.

In this paper, I will criticize this “explanatory argument” for the identity of phenomenal and physical properties. My starting point is Kim’s (2005) attack on the explanatory argument, and I will focus on McLaughlin’s (2007, 2010) and Bates’ (2009) formulation of the argument. What I will show is that neither Kim nor the supporters of the explanatory argument have taken sufficiently into account the actual role that identities and correlations play in scientific explanations. I will argue that (*contra* Kim) identities do play a crucial role in scientific explanation, but (*contra* the new type physicalists) they are not put forward as explanations for correlations of the kind that are involved in the case of phenomenal properties. For this reason, a crucial premise in the explanatory argument for type physicalism is undermined by scientific practice, and hence the argument fails to provide new grounds for accepting type physicalism.

## **2. The explanatory argument for type physicalism**

A state of phenomenal consciousness is such that there is something that it is like to be in that state. For example, there is something it is like to be in pain, or see red, or smell sulphur. Phenomenal states<sup>1</sup> are notoriously resistant to explanations in physical or neural terms, as several passionately debated arguments purport to show (the explanatory gap, Nagel’s bat, Mary in a black-and-white room, and so on). However, it is plausible that

some day scientists will have discovered exactly how phenomenal states are *correlated* with neural or physical states. This can be formulated as the following *correlation thesis*: “For any type of state of phenomenal consciousness C there is a type of physical state P such that it is true and counterfactual supporting that a being is in C if and only if the being is in P” (McLaughlin, 2010, p. 267).

The truth of the correlation thesis would not yet settle the question of the place of phenomenal consciousness in nature, since the thesis is compatible with various different ontological positions, including property dualism, neutral monism, parallelism and type physicalism. These positions each have their problems and virtues, which have been widely discussed. However, the novel strategy of the supporters of the explanatory argument is to take the truth of the correlation thesis as a starting point and then argue that we are justified in believing that type physicalism is true because it *explains* the correlation thesis, while the other ontological solutions leave it unexplained. Type physicalism states that (*Type Physicalism*): “For every type of state of phenomenal consciousness C, there is a type of physical state P such that  $C = P$ ” (McLaughlin, 2010, p. 266). This indeed seems to provide a straightforward explanation for the correlation thesis: for each state of phenomenal consciousness, the reason why a being is in C if and only if the being is in P is that  $C = P$ . In contrast, for example, (emergent) property dualism simply asserts that psychoneural correlations are a fundamental unexplainable feature of the world.

It seems then that Type Physicalism is the *best explanation* for the correlation thesis. This follows immediately if we accept that Type Physicalism explains the correlation thesis while all the alternative ontological solutions do not explain it at all (or fail for some other reason). Finally, supporters of the argument claim that the fact that the identity thesis is the best explanation for the correlation thesis provides the *justification* for holding the identity thesis to be true. The idea behind this is that if we know that A is true and that B is the best explanation for A, then we are justified in believing B.

The explanatory argument can also be presented more schematically as follows (adapted from Bates, 2009, p. 315):

**P1. *Correlation Thesis*:** For any type of state of phenomenal consciousness C there is a type of physical state P such that it is true and counterfactual-supporting that a being is in C if and only if the being is in P.

**P2.** One possible explanation of the Correlation Thesis is *Type Physicalism*: For every type of state of phenomenal consciousness C, there is a type of physical state P such that  $C = P$ .

**P3.** No alternative explanation of the Correlation Thesis is as good as Type Physicalism.

**Conclusion:** Type Physicalism is true.

### 3. Kim's Objection

Jaegwon Kim (2005, ch. 5) has presented several objections to the explanatory argument. For example, Kim argues that the very idea of correlation implies that the correlates are distinct, and raises doubts about the principle of inference to the best explanation. However, I will only discuss here the objection that I consider to be the strongest one, and which also provides the starting point for my own criticism. This objection states that psychophysical identities do not explain psychophysical correlations, or anything else for that matter, and hence the explanatory argument fails. According to Kim (2005), scientific explanations of correlations are nothing like McLaughlin and others take them to be:

In science there seem to be two principal ways of explaining correlations: first, correlations are sometimes explained by invoking a single lower-level process or

structure underlying the correlated phenomena; second, the explanation may proceed by showing the correlated phenomena to be collateral effects of a common cause. ... In any case, it is quite obvious that scientists will not in general attempt to explain correlations by identifying the correlated properties. (p. 134)

Before discussing these claims, it is important to note that the sense in which ‘correlation’ is used in the explanatory argument is very different from the standard scientific use of the term, which is statistical correlation. This is something Kim is not clear enough about: talk about scientific explanations of correlations immediately brings into mind statistical correlations. However, the correlation thesis only says: “For any type of state of phenomenal consciousness *C* there is a type of physical state *P* such that it is true and counterfactual supporting that a being is in *C* if and only if the being is in *P*”. As McLaughlin (2010, p. 271) points out, it does not even involve the term correlation, except in the title. Therefore, it would be more accurate to talk of nomological copresence instead of correlation. In the rest of this paper, I mean by ‘correlation’ the kind of nomological copresence expressed by the correlation thesis.

In any case, the type physicalists’ answer to this point is straightforward. They argue that there is a third way of explaining correlations in science, and that is by pointing out that the correlates are identical. According to McLaughlin (2010, p. 282), “ $A = B$ ” can explain why *A* is present when and only when *B* is present. To support this claim, he provides an example: when Maxwell realized that the speed of electromagnetic waves in a vacuum is the speed of light, he made the “bold conjecture” that “light waves = electromagnetic waves.” McLaughlin argues that the identity “light waves = electromagnetic waves” explains why light waves are present when and only when electromagnetic waves are present.

I agree with the type physicalists that Kim has not convincingly shown that identities do not play *any role* in scientific explanations. Scientific hypotheses often involve identity claims, and these claims are explanatorily substantial. However, as I will show in the next section, I do think that Kim's basic intuition is right: in science, identities are *not* put forward to explain correlations of the kind that the correlation thesis involves. For this reason, the explanatory argument fails.

#### **4. The Role of Identities in Scientific Explanation**

The discussion on type physicalism in philosophy of mind has traditionally focused on conceptual and metaphysical issues related to the nature or possibility of psychoneural identities. Little attention has been paid to science and the actual role identities play in scientific explanations. One of the positive features of the debate on the explanatory argument is that it is at least superficially connected to science, and both proponents and opponents of the argument appeal to science in support of their views, as we also saw in the discussion between McLaughlin and Kim above. Particularly notable is that the explanatory argument is presented as an (partly) empirical argument for type physicalism. McLaughlin explicitly claims that Type Physicalism is an empirical hypothesis, and that it figures in the explanation of the correlation thesis in a similar way as hypotheses in science figure in explanations (McLaughlin, 2010, p. 279; see also Bates, 2009, p. 320, and Melnyk, 2003, p. 9-10). However, the evidence from science presented in support of this is rather scarce. McLaughlin mainly appeals to the case that already came up in the previous section: the identity of light waves and electromagnetic waves. In more detail, McLaughlin (2010) writes:

When Maxwell's calculations showed that electromagnetic waves have the same speed in a vacuum as the known speed of light, he famously made "the bold conjecture" that light waves = electromagnetic waves ... The hypothesis that light waves are electro-magnetic waves was invoked to explain why (1) electromagnetic waves and light waves occur in the same spatial regions at the same time, why (2) electromagnetic waves have the same speed in a vacuum as light waves, and why (3) the refractive indices in materials are exactly the same for light waves and electro-magnetic waves. (p. 282)

I grant that this example shows that (*contra* Kim) identities do play a role in scientific explanation. However, it does not show that identities are a way of explaining *correlations*, understood as nomological copresence. The crucial part in Maxwell's argument was showing that electromagnetic waves and light waves propagate at the same velocity (speed of light). After considering the velocity of propagation of "magnetic disturbances", Maxwell concluded that "[t]his velocity is so nearly that of light, that it seems that we have strong reason to conclude that light itself (including radiant heat, and other radiations if any) is an electromagnetic disturbance in the form of waves propagated through the electromagnetic field according to electromagnetic laws" (Maxwell, 1865, p. 466).

Importantly, the fact that light waves are present when and only when electromagnetic waves are present did not play a role in making (or justifying) the "bold hypothesis". The identity claim was not presented to explain this correlation, and such a correlation alone would never have been enough to justify the bold hypothesis. It was presented to explain the fact that light waves and electromagnetic waves share an important property (the speed of propagation). Another crucial point is that Maxwell's claim was a hypothesis that still had to be confirmed by further *empirical* tests and evidence. In this case, the empirical confirmation came relatively soon, as Heinrich Hertz

discovered in the 1880s that electromagnetic waves exhibit certain crucial properties of light waves, such as refraction, reflection, and interference.<sup>2</sup>

To get a better picture of the actual role that identities play in scientific explanations, let us consider another example from a field somewhat closer to philosophy of mind, namely vision research. About 40 years ago, Barlow and Levick were studying the responses of ganglion cells in the cat retina at different luminance levels (Barlow & Levick, 1969). They discovered strangely behaving “luminance units” that were extremely rare (less than 1% of the retinal ganglion cell population) and responded to light stimuli in an unusual way: the response was sluggish, relatively straightforwardly related to the light intensity, and increased monotonically with increasing light intensity (normally the responses of retinal ganglion cells are much more complex).

In a different line of research, a population of morphologically distinct retinal ganglion cells containing the photopigment melanopsin has recently been discovered in the mammalian retina (see Do & Yau (2010) for a review). Interestingly, these melanopsin-containing cells are intrinsically photosensitive – they respond to light even when the synaptic transmission from the normal photoreceptors (rods and cones) is blocked. They resemble the “luminance units” in several important ways: their response to light stimuli is sluggish, the relation between response and light intensity is monotonic, and they are also very rare (comprising about 1-3 % of the retinal ganglion cell population).

Based on these common properties, it seems extremely likely that the melanopsin-containing cells *are* the luminance units. However, this identity remains hypothetical, since it has not (yet) been shown that all the properties of luminance units are also exhibited by the melanopsin-containing cells (or the other way around) (Do & Yau, 2010). Thus, this is a case of a hypothetical identity claim that is pending empirical confirmation.



These two examples (electromagnetic waves = light waves, melanopsin containing cells = luminance units) nicely illustrate how identities figure in scientific explanations. Other examples that exhibit the same pattern are easy to name. For example, rod cells in the retina keep signaling in fully dark conditions, and the level of this “dark light” was measured in the 1940s and 1950s in human psychophysical experiments. Later the level of quantal noise created by the photopigment (rhodopsin) molecules in the rod cells was measured, and it turned out to match exactly the level of dark light measured psychophysically.<sup>3</sup> This led to the very plausible hypothesis that the quantal noise *is* the dark light (Schneeweis & Schnapf, 2000). One more example of this pattern would be the (controversial) hypothesis that the mechanism of long term potentiation (LTP) is the mechanism for memory consolidation (see, e.g., McGaugh, 2000).

Two characteristic features of these identity claims are essential in the present context. First, they are put forward as hypotheses to explain why two things that were initially believed to be distinct, or were discovered by different methods, both have the same or similar properties. The fact that light waves and electromagnetic waves propagate at the same speed and have the same refractive indices is explained by the hypothesis that light waves are electromagnetic waves. The fact that luminance units and melanopsin-containing ganglion cells react to light in the same way is explained by the hypothesis that luminance units are melanopsin-containing ganglion cells, and so on.

The second feature is even more important: hypothetical identities are to be tested just like any other hypotheses in science, by seeing how they stand up to empirical evidence. One way of doing this is by testing the further hypotheses that the identity claim suggests. Since we know that if two things are identical, they have to have exactly the same properties (*indiscernibility of identicals*), a hypothetical identity claim immediately provides ways of testing itself: if A is identical with B and A has property P, then B also has to have property P, and so on. This is also how Maxwell’s hypothesis was

tested and confirmed: if light waves are electromagnetic waves, then both must have the same refractive indices, and this indeed turned out to be the case.

This second point about the role of scientific identity claims is strongly supported by the “Heuristic Identity Theory” (HIT) approach (Burnston, Sheredos, & Bechtel, 2011; McCauley, 2012; McCauley & Bechtel, 2001; the theory is largely based on the insights of William Wimsatt (1976)). The proponents of this theory argue that identities are not conclusions of scientific work but hypothetical premises – identity claims are the *beginning* of scientific inquiry. The principal motivation for formulating identity claims is their potential to advance empirical research. In particular, the role of identities is *not* to explain correlations, but to connect levels or sciences and to generate new hypotheses and new avenues for research. This is one crucial aspect in which identities differ from correlations: they suggest explanatory connections that demand further empirical exploration, while correlations do not. In this sense, claims about correlations and claims about identities are “different conceptual animals that thrive in different theoretical habitats” (McCauley and Bechtel, 2001, p. 754).

## **5. Type Physicalism as a Hypothetical Identity**

We are now in a position to see how the identity claim that the proponents of the explanatory argument defend differs from these scientific identity claims. The type physicalists’ claim is that “for every type of state of phenomenal consciousness C, there is a type of physical state P such that  $C = P$ ”. First of all, this claim differs from the scientific identity claims because it is not posited to explain any shared properties; it is posited to explain the nomological copresence of physical and phenomenal properties. Secondly, it differs from them because it cannot be empirically tested in the way scientific identity claims are – the claim does *not* lead to any predictions or further

hypotheses that could be empirically verified. Proponents of the explanatory argument would certainly agree that there is no further empirical evidence or tests that could distinguish between Type Physicalism and the other ontological positions – their claim is that Type Physicalism is superior based “holistic grounds of overall coherence and simplicity with respect to total theory” (McLaughlin, 2007, p. 436). If there was an empirical test or a prediction that could distinguish between the ontological alternatives, all this philosophical discussion would be unnecessary: we could just wait until the empirical results are in.<sup>4</sup>

Hence, the identity claim of the type physicalists fundamentally differs from the scientific identity claims that I have discussed in the previous section. These considerations also point to a further problem in the explanatory argument for type physicalism. In all of the examples discussed, the hypothetical identity is an identity of *things*. This includes also most of the examples of successful identifications that McLaughlin (2010) and Bates (2009) present (Bill Sikes = the burglar, water = H<sub>2</sub>O, Tully = Cicero, light waves = electromagnetic waves, etc.). However, what is at issue in the case of psychoneural correlations and type physicalism is the identity of *properties*. The claim is that the correlation of properties is explained by the fact that the properties are identical.

To be fair, the proponents of the explanatory argument do discuss one case that clearly involves properties – the case of mean kinetic energy and heat (e.g., Block & Stalnaker, 1999, p. 23-24). Furthermore, perhaps the other cases can be translated into talk of properties as well, for example “being water = being H<sub>2</sub>O”. However, the disanalogy between cases such as these and the case of phenomenal properties is well known and has been thoroughly discussed in philosophy of mind (Levine (1983) is one *locus classicus*). The identity of mean kinetic energy and heat is justified by the fact that mean kinetic energy plays the same causal role as heat, while phenomenal properties and physical properties do not have the same causal role (phenomenal properties resist a

causal analysis). Therefore, these are certainly not examples where the identity claim is put forward to explain the nomological copresence of the properties.

The supporters of the explanatory argument would need to provide examples from science where the correlation of two properties is explained by these properties being identical, but such examples are hard to find. It is not a common explanatory strategy in science to explain the copresence of properties by hypothesizing that they are identical. In fact, it is generally the case that correlated properties are *not* identical (see also Towl, 2012, who discusses the different ways neuroscientists can explain correlations *without* appealing to identities). Consider for example the property of having a heart and the property of having a circulatory system. An organism has a heart if and only if it also has a circulatory system. However, explaining this correlation with the hypothetical identity “the property of having a heart = the property of having a circulatory system” would be deeply mistaken.<sup>5</sup> Another everyday example of nomologically copresent properties that are not identical is being a three-dimensional object and having a volume: the property of being a three-dimensional object is present when and only when the property of having a volume is present. To take a physical example, having a half-integer spin is nomologically copresent with being a fermion, but having a half-integer spin is not identical to being a fermion. It seems that when two properties are systematically correlated, but do not have anything else in common (such as their causal role), there are good reasons to assume that the properties in question are *not* identical.

One could argue that phenomenal properties and physical properties do share one property, namely their spatio-temporal location. However, this does not bring the case any closer to the scientific cases I have discussed in the previous section. Even if we accept that spatio-temporal location is a property, the main point holds: scientists do not posit identities merely to explain shared spatio-temporal location. This is supported by the several examples mentioned above, where properties that have the same spatio-temporal location and are nomologically copresent are not identical (e.g., the property of

being a fermion and the property of having a half-integer spin). If spatio-temporal location counts as a property, we merely have to reformulate the claim: if two properties are systematically correlated and the only property they share is their spatio-temporal location, there are good reasons to assume that the properties are not identical.

These considerations cast serious doubts on the central premise of the explanatory argument, namely the assumption that the identity thesis (Type Physicalism) provides the best explanation for the correlation thesis. If we make the plausible assumption that identity claims in philosophy should at least remotely resemble identity claims in science, or that the requirements for an explanatory identity claim are similar in philosophy as in science, then it seems that Type Physicalism is not an explanatory identity claim. Since the type physicalists have not provided a single example from science where the copresence of properties is explained by those properties being identical, it is not at all clear that one possible scientific explanation for the correlation thesis is Type Physicalism. In any case, the burden of proof is on the supporters of the explanatory argument to show why we should consider Type Physicalism to be a satisfactory explanation for the nomological copresence of phenomenal and physical properties.

In response to this, type physicalists might claim that there is a straightforward and intuitively clear sense in which identities can explain correlations. For example, it does seem intuitively quite plausible that the fact that water is present when and only when  $H_2O$  is present is explained by the hypothesis that water is  $H_2O$  (see also McLaughlin, 2010, p. 283-284). However, when considering intuitions like this, it is crucial to distinguish between two epistemic situations: (1) the epistemic situation where we already know that water is  $H_2O$ , and (2) the imaginary epistemic situation where we do not yet know that water is  $H_2O$ ; we know only that water is nomologically copresent with  $H_2O$ . In the first situation, it is not clear whether the idea of explaining the copresence of water and  $H_2O$  even makes sense – the question why water and  $H_2O$  are copresent will never arise. In any case, the explanation will at best be shallow and trivial;

it will not increase our understanding of the world. In the second situation, which resembles the case of phenomenal and physical properties, we are at best warranted to make the *hypothesis* that water is H<sub>2</sub>O, and this hypothesis needs to be confirmed by further research. The hypothesis is only the *starting point* for inquiry into the nature of water and H<sub>2</sub>O. It explains the correlation between water and H<sub>2</sub>O, if at all, in a shallow and preliminary way. Therefore, examples or intuition pumps like these do not provide support for the explanatory argument.

One might try to move the discussion to a more general level and claim that, in spite of all that has been said, the psychophysical correlations do count as scientific evidence in support of Type Physicalism. This may well be true, but it does not undermine my criticism in any way. The explanatory argument requires that Type Physicalism is the *best explanation* the correlation thesis, and this is something very different from psychophysical correlations conferring evidence for Type Physicalism. One might also claim that Type Physicalism is the philosophically least problematic ontological solution to the problem of phenomenal properties. Again, this may very well be true, but it is beside the point, since the explanatory argument is supposed to be a *new* and *empirical* argument for Type Physicalism.

It is also important to note that the proponents of HIT explicitly apply their theory to the case of consciousness, arguing that hypothesizing identities between conscious states and neural states is a reasonable research strategy that leads towards a kind of type identity theory (McCauley and Bechtel, 2001, p. 752-757). However, this approach is very different from the best explanation strategy, since McCauley and Bechtel do not accept the correlation thesis as a starting point or as something that is in need of explanation: they claim that construing neuroscientific accounts of consciousness in terms of correlations is a misguided approach. Therefore, the main argument of this paper does not apply to the HIT account of consciousness.<sup>6</sup>

In general, it is important to emphasize that I do not claim to have refuted Type Physicalism. In fact, I have presented no argument against it. The only aim of this paper is to point out a crucial shortcoming in one recently popular argument for Type Physicalism, namely the explanatory argument.

## **6. Concluding remarks**

In this paper, I have argued that the explanatory argument for Type Physicalism is not supported by the role of identity claims in science. Scientists do present identities as explanatory hypotheses, but not in a way that the explanatory argument requires. In science, an identity claim is typically the starting point for further research, and motivated by shared properties, not mere correlations. The supporters of the explanatory argument have so far failed to show why and how Type Physicalism would provide a scientific explanation for the correlations between phenomenal and physical properties.

In this light, the premise **P2** (“One possible explanation for the correlation thesis is Type Physicalism”) of the explanatory argument can be questioned. Furthermore, even if we grant that this premise holds, the considerations in this paper cast serious doubts on the explanatory argument. The core of the argument is that we are justified to believe that Type Physicalism is true for the reason that no alternative explanation of the psychoneural correlations is as good as Type Physicalism. An inference to the best explanation of this kind is not at all straightforward or unproblematic. As Lipton (2000) points out, “inference to the best explanation” must generally be glossed as “inference to the best of the available competing explanations, when the best one is sufficiently good”. In the present case, the question whether the best explanation is sufficiently good is particularly pressing, considering the issues I have raised above. If even the best

explanation available is problematic or shallow, which seems to be the situation in this case, then an inference to the best explanation is not warranted.

To conclude: the burden of proof is on the type physicalists to show how the identity thesis explains the correlation thesis, and why this explanation is sufficiently good to warrant an inference to the best explanation. As it stands, the explanatory argument gives us no new empirical grounds to accept Type Physicalism.

## Notes

1. Later in the paper I mostly talk of properties instead of states. For example, phenomenal properties can be seen as the qualitative characters of phenomenal states. Nothing substantial turns on this distinction.

2. The fact that Maxwell's theory (including the identity of light waves and electromagnetic waves) was still considered a hypothesis that was in need of confirmation, and that Hertz himself saw his work as exactly this kind of confirmation, is nicely illustrated by the following quote (Hertz, 1893, p. 19-20): "From the outset Maxwell's theory excelled all others in elegance and in the abundance of the relations between the various phenomena which it included. The probability of this theory, and therefore the number of its adherents, increased from year to year. But as long as Maxwell's theory depended solely upon the probability of its results, and not on the certainty of its hypotheses, it could not completely displace the theories which were opposed to it. The fundamental hypotheses of Maxwell's theory contradicted the usual views, and did not rest upon the evidence of decisive experiments. In this connection we can best characterize the object and the result of our experiments by saying: The object of



these experiments was to test the fundamental hypotheses of the Faraday-Maxwell theory, and the result of the experiments is to confirm the fundamental hypotheses of the theory.”

3. Interestingly, in all of these cases, what initially motivates the hypothetical identity is the similarity of a certain quantitatively measurable property (speed of propagation, pattern of response, or level of signaling). The likelihood that the hypothesis is true then partly depends on how precisely this property can be measured. If the quantities match exactly when measured precisely, this is very strong evidence in favor of the hypothetical identity. In the case of the speed of propagation of light waves and electromagnetic waves, the measurements that Maxwell had at his disposal had large margins of error, which was one reason why the hypothesis needed further testing. Similarly, in the case of luminance units and melanopsin-containing ganglion cells, the response patterns have at present time not yet been measured or quantified precisely enough to confirm the hypothetical identity (Do & Yau, 2010, p. 1552).

4. It is important to remember that the issue concerns *phenomenal* properties, which resist causal analyses. There may very well be ways of empirically testing various ontological hypotheses regarding *non-phenomenal* mental properties, which can be causally or functionally defined.

5. Admittedly, it is not entirely clear whether it is counterfactual supporting that the property of having a heart is present when and only when the property of having a circulatory system is present. This depends on how exactly the example is construed and what kinds of counterfactuals need to be considered. However, the other examples (being three-dimensional  $\leftrightarrow$  having a volume, being a fermion  $\leftrightarrow$  having a half integer spin) are clearly counterfactual supporting.

6. However, when it comes to phenomenal consciousness, some of the issues I have raised also concern the HIT approach. I have pointed out that the case of phenomenal consciousness does not resemble the situations where hypothetical identities are posited

in science. What then justifies the hypothetical identity of phenomenal and neural properties? What are the new hypotheses and avenues of research that this identity discloses? What are the further tests that could confirm or disconfirm the identity claim? Perhaps the problem lies in the way phenomenal consciousness and phenomenal properties have been traditionally construed in this debate, which is something supporters of HIT need not accept. In any case, these points clearly call for a more detailed discussion that goes beyond the scope of this paper.

### **Acknowledgements**

I am very thankful for helpful comments that several people have offered on earlier versions of this paper. These people include, in no particular order: Achim Stephan, Brian McLaughlin, Dan Brooks, Raphael van Riel, Leon de Bruin, Albert Newen, Bob Richardson, Laura Bringmann, and the two anonymous referees of this journal. A special thank goes to Malte Ahlers for helping me with the details of retina research. I have also presented different versions of this paper at the ECAP7 in Milan and at an internal meeting of our research group in Bochum, receiving very useful feedback. This work was made financially possible by generous grants from the Finnish Cultural Foundation, HWK Institute for Advanced Study, and Ruhr University Bochum.

### **References**

Barlow, H. B., and Levick, W. R. (1969). Changes in the maintained discharge with adaptation level in the cat retina. *Journal of Physiology*, 202, 699-718.

- Bates, J. (2009). A defence of the explanatory argument for physicalism. *The Philosophical Quarterly*, 59, 315-324.
- Block, N. & Stalnaker, R. (1999). Conceptual analysis, dualism, and the explanatory gap. *The Philosophical Review*, 108, 1-46.
- Burnston, D., Sheredos, B., & Bechtel, W. (2011). HIT on the Psychometric Approach. *Psychological Inquiry* 22, 108-114.
- Do, M. T. H., & Yau, K.-W. (2010). Intrinsically Photosensitive Retinal Ganglion Cells. *Physiological Reviews*, 90, 1547-1581.
- Hertz, H. R. (1893). *Electric waves: being researches on the propagation of electric action with finite velocity through space*. (D. E. Jones, Transl.). London: MacMillan.
- Hill, C. S. (1991). *Sensations: A Defense of Type Materialism*. Cambridge: Cambridge University Press.
- Kim, J. (2005). *Physicalism, or Something Near Enough*. Princeton: Princeton University Press.
- Levine, J. (1983). Materialism and qualia: the explanatory gap. *Pacific Philosophical Quarterly*, 64, 354-361.
- Lipton, J. (2000). Inference to the Best Explanation. In W. H. Newton-Smith (ed.) *A Companion to the Philosophy of Science* (pp. 184-193). Malden, MA: Blackwell.
- Maxwell, J. C. (1865). A Dynamical Theory of the Electromagnetic Field. *Philosophical Transactions of the Royal Society of London* 155, 459-512.  
DOI:10.1098/rstl.1865.0008
- McCauley, R. N. (2012). About Face: Philosophical Naturalism, The Heuristic Identity Theory, and Recent Findings about Prosopagnosia, In S. Gozzano & C. S. Hill (eds.) *New Perspectives on Type Identity: The Mental and the Physical* (pp. 186-206). Cambridge: Cambridge University Press.

- McCauley, R. N. & Bechtel, W. (2001). Explanatory pluralism and heuristic identity theory. *Theory & Psychology*, 11, 736-760.
- McGaugh, J. L. (2000). Memory – a Century of Consolidation. *Science* 287, 248-251.
- McLaughlin, B. P. (2001). In Defense of New Wave Materialism: A Response to Horgan and Tienson. In C. Gillett & B. Loewer (eds.) *Physicalism and Its Discontents* (pp. 319-330). Cambridge: Cambridge University Press.
- McLaughlin, B. P. (2007). Type materialism for phenomenal consciousness. In M. Velmans & S. Schneider (eds.) *The Blackwell Companion to Consciousness* (pp. 431-444). Malden, MA: Blackwell.
- McLaughlin, B. P. (2010). Consciousness, type physicalism, and inference to the best explanation. *Philosophical Issues*, 20, 266-304.
- Melnyk, A. (2003). *A Physicalist Manifesto: Thoroughly Modern Materialism*. Cambridge: Cambridge University Press.
- Schneeweis, D. M., & Schnapf, J. L. (2000). Noise and light adaptation in rods of the macaque monkey. *Visual Neuroscience*, 17, 659-666.
- Towl, B. N. (2012). Mind-brain correlations, identity, and neuroscience. *Philosophical Psychology*, 25, 187-202.
- Wimsatt, W. C. (1976). Reductionism, levels of organization, and the mind-body problem. In Globus et al. (eds.) *Consciousness and the Brain. A Scientific and Philosophical Inquiry* (pp. 205-267). New York: Plenum Press.